Attributes and Rough Properties in Information Systems

Keh-Hsun Chen and Zbigniew W. Ras
Department of Computer Science, University of North Carolina
Andrzej Skowron
University of Warsaw, Institute of Mathematics

ABSTRACT

We introduce the notion of an artificial attribute in Scott's systems. Assuming that only some attributes are visible, we build an approximation space in Schott's system and then we talk about rough properties in Pawlak's sense. A first-order language with a standard interpretation in the mentioned approximation space is the main tool in our investigations.

KEYWORDS: ***approximation space, artificial attributes, first-order language, information system, interpretation, rough properties.***

INTRODUCTION

There are a number of algebraic models of information systems. They have been proposed by Salton [1], Pawlak [2], Scott [3], and others. In Salton's model the set of descriptions is partially ordered and the classification function, which is a function from the set of descriptions into the power set of objects, is monotonic. In Pawlak's model the set of descriptions is built from the values of attributes. These attributes have a strong influence on classification methods. Ras [4] proposed the notion of an artificial attribute in Salton's system to show some similarities between these two models.

The starting point in our investigation is the information system proposed by Scott [3]. Objects are represented as consistent, usually infinite "closed" sets of propositions. Descriptions (elements of Con) are represented as finite consistent sets of propositions. We use the definition of Con as a departure point for getting

Address correspondence to Professor Z. W. Ras, Department of Computer Science, University of North Carolina, Charlotte, North Carolina 28223.

International Journal of Approximate Reasoning 1988; 2:365–376

52 Vanderbilt Ave., New York, NY 10017 0888-613X/88/$3.50

into the structure of Scott's system more deeply. We propose a notion of an artificial attribute in Scott's system that has all the nice properties that attributes have in complete information systems: their values describe disjoint sets of objects and cover the whole set of objects.

Scott's information system is infinite; more precisely, it has an infinite set of propositions. We have proved that if an infinite information system has only finite artificial attributes, then the number of its artificial attributes is infinite. From the practical point of view, only a finite set of propositions are visible (in other words, only a finite set of attributes are visible). This situation forced us to define an approximation space (see Pawlak [2]) generated by visible attributes. Having this approximation space we may now talk about rough properties (Pawlak [2, 5, 6]) in Scott's system. To do that we consider a first-order language L. We deal only with interpretations of L in the mentioned approximation space. We look for the lower and upper approximations for any formula and term of L.

Information Systems and PO-Systems

In this section, we recall the definitions of information system (see Scott [3]) and PO-system (see Ras [4]) and show that an information system can be viewed as a PO-system. This opens the door to artificial attributes of information systems.

By an information system we mean a structure $A = (D, \Delta, \text{Con}, \vdash)$ where D is a set of propositions, $\Delta \in D$ is the least informative proposition, Con is a set of finite subsets of D (the finite consistent sets of propositions), and $\vdash$ is a binary relation between members of Con and members of D (the entailment relation for objects). The following axioms must be satisfied:

(i) If $u \in \text{Con}$ and $v \subseteq u$, then $v \in \text{Con}$.
(ii) If $X \in D$, then $\{X\} \in \text{Con}$.
(iii) If $u \vdash X$, then $u \cup \{X\} \in \text{Con}$.
(iv) $u \vdash \Delta$
(v) $u \vdash X$ for any $X \in u$
(vi) If $(\forall Y \in u)(v \vdash Y)$ and $u \vdash X$, then $v \vdash X$.

Let us extend the relation $\vdash$ onto $\text{Con} \times P(D)$ as follows:

$$u \vdash v \equiv (\forall X \in v)[u \vdash X]$$

By $P(D)$ we denote the power set of D. If $v = \{X\}$, then $u \vdash v$ is equivalent to $u \vdash X$. Condition (vi) in the above definition will be replaced by

(vi′) If $v \vdash u$ and $u \vdash w$, then $v \vdash w$ for any $v, u \in \text{Con}$, $w \in P(D)$.

The elements of $A = (D, \Delta, \text{Con}, \vdash)$ are those subsets x of D where

1. All finite subsets of x are in Con
2. If $u \subseteq x$ and $u \vdash Y$, then $Y \in x$

By $|A|$ we denote the set of all elements of A. Clearly $(|A|, \subseteq)$, where $\subseteq$ is the set-theoretical inclusion, is a partially ordered set. By Tot_A we denote the set of all maximal elements in $(|A|, \subseteq)$. The elements of Tot_A are said to be total.

Let $\bar{u} = \{X \in D{:}u \vdash X\}$ for any $u \in \mathrm{Con}$. Clearly $\bar{u} \in |A|$.

For any $x \in |A|$ we have the basic formula:

$$x = \bigcup \{\bar{u} : u \in \mathrm{Con} \text{ and } u \subseteq x\}$$

Let us now take the information function

$$\delta : P(D) \rightarrow P(\mathrm{Tot}_A)$$

where $\delta(z) = \{x \in \mathrm{Tot}_A{:}z \subseteq x\}$.

FACT 1

(i) If $y \subseteq x$ then $\delta(x) \subseteq \delta(y)$
(ii) $\delta(x \cup y) = \delta(x) \cap \delta(y)$
(iii) $\delta(x \cap y) \supseteq \delta(x) \cup \delta(y)$
(iv) $\delta(\bigcup_i x_i) = \bigcap_i \delta(x_i)$

FACT 2 If x is finite then $\delta(x) = \varnothing$ iff $x \notin \mathrm{Con}$.

Proof If $y \in \delta(x)$ then $x \subseteq y$ and $y \in |A|$. Hence $x \in \mathrm{Con}$. If $x \in \mathrm{Con}$ then $x \subseteq \bar{x}$, and $\bar{x}$ can be extended to a total element.

Observe that $S = (\mathrm{Tot}_A, \mathrm{Con}, \subseteq, \delta)$ is a PO-system according to the definition given by Ras [4]. In this case Tot_A is interpreted as the set of objects, Con as the set of requests, and $\delta(z)$ as the set of objects retrieved by the request $z \in \mathrm{Con}$.

The definition of a PO-system requires that (Con, $\subseteq$) has to be a po-set, δ has to be a monotonic function, and $(\forall z \in \mathrm{Tot}_A)(\exists x \in \mathrm{Con})(z \in \delta(x))$. In our case all these requirements are satisfied.

Artificial Attributes in Information Systems

In this section (artificial) attributes of information systems are developed and studied. A notion of an artificial attribute for a PO-system was defined by Ras [4]. To apply the main idea of that notion here, we define a subset of Con containing requests retrieving subsets of Tot_A that form a partition of Tot_A.

Let us observe that the structure of (Con, $\subseteq$) is determined by the six conditions required by Con. This means that the structure of (Con, $\subseteq$) is not as arbitrary as in the case of general PO-systems. This allows us to look for artificial attributes for S in a somewhat different way than in Ref 4.

DEFINITION Let $u, v \in \mathrm{Con}$. We say that u and v are contradictory if $u \cup v \notin \mathrm{Con}$.

Assume that B is a maximal subset of Con such that any two elements in it are contradictory. The question arises whether

$$\bigcup \{\delta(u) : u \in B\} = \mathrm{Tot}_A$$

FACT 3 If $x \in \mathrm{Tot}_A$ is finite then $(\exists y \in B)(y \subseteq x)$.

Proof Since x is finite, $x \in$ Con. Assume $\sim(\exists y \in B)(y \subseteq x)$. Let's consider $B \cup \{x\}$. Either (i) $x \cup y \in$ Con for some $y \in B$ or (ii) $x \cup y \notin$ Con for all $y \in B$. In case (i) $x \nsubseteq x \cup y \in$ Con, which implies $x \notin \mathrm{Tot}_A$. In case (ii) $B \nsubseteq B \cup \{x\}$, B cannot be maximal.

FACT 4 If $x \in \mathrm{Tot}_A$ is denumerable and B is finite then $(\exists y \in B)(y \subseteq x)$.

Proof Assume that N denotes the set of positive integers, $x = \bigcup_{i \in N} x_i$, where $x_i \in$ Con, $x_i \subseteq x_{i+1}$ for $i \in N$. Assume moreover that

$$\sim(\exists y \in B)(y \subseteq x)$$

Following the previous proof, let's take $B \cup \{X_i\}$, where $i \in N$. We have two options:

(i) $x_i \cup y \notin$ Con for any $y \in B$.
(ii) $x_i \cup y \in$ Con for some $y \in B$.

If there is an $i \in N$ satisfying (i), then B is not maximal. So we assume that

$$(\forall i \in N)(\exists y(i) \in B)(x_i \cup y(i) \in \mathrm{Con})$$

Since B is finite, there exists an element $y \in B$ and an infinite, increasing sequence $(i_1, i_2, \ldots)$ of elements from N such that

$$y = y(i_j) \text{ for all } j \geqslant 1$$

Hence

$$\bigcup_{j \geqslant 1} (x_{i_j} \cup y(i_j)) = \bigcup_{j \geqslant 1} x_{i_j} \cup y = x \cup y$$

Clearly any finite subset of $x \cup y$ is in Con, which implies that $x \notin \mathrm{Tot}_A$.

DEFINITION By an artificial attribute in $S = (\mathrm{Tot}_A, \mathrm{Con}, \subseteq, \delta)$ we mean any maximal subset $B \subseteq$ Con such that

$$(\forall u \in B)(\forall v \in B)(u \neq v \rightarrow u \cup v \notin \mathrm{Con})$$

THEOREM 1 If D is a denumerable set and $B = \{b_1, \ldots, b_n\}$ is an artificial attribute in S, then $\{\delta(b_i)\}_{i=1}^{n}$ is a partition of Tot_A.

Proof If D is denumerable, then any $x \in \mathrm{Tot}_A$ is denumerable. Applying Fact 4 we have that $\{\delta(b_i)\}_{i=1}^{n}$ covers Tot_A. From Fact 1 we have that

$$\delta(b_i) \cap \delta(b_j) = \delta(b_i \cup b_j)$$

where $b_i \cup b_j \notin$ Con. Now applying Fact 2 we have that $\delta(b_i \cup b_j) = \varnothing$.

Below we given an example that shows that the assumption saying that B is finite is essential.

EXAMPLE Let $A = (D, \Delta, \mathrm{Con}, \vdash)$ be an information system where

$$D=\{n=0, n=1, n=2, \cdots\} \cup \{n \geqslant 0, n \geqslant 1, n \geqslant 2, \cdots\}$$

A subset of D is consistent if it does not contain contradictory statements. For instance, $\{n = 1, n \geqslant 2\}$ is not consistent but $\{n = 3, n \geqslant 2\}$ is consistent. Having defined Con, let us observe that

$$\{\{n=0\}, \{n=1\}, \{n=2\}, \cdots\}$$

is a maximal subset of Con such that any two elements in it are contradictory. Clearly, for total element

$$x=\{n \geqslant 0, n \geqslant 1, n \geqslant 2, \cdots\}$$

the inclusion $\{n = i\} \subseteq x$ does not hold for any $i \geqslant 0$.

THEOREM 2 Let $A = (D, \Delta, \mathrm{Con}, \vdash)$. For every $a \in D$ there exists an artificial attribute B such that $\{a\} \in B$.

Proof Let $F = \{C{:}\{a\} \in C$ and C is a family of pairwise contradictory sets in Con$\}$. Then F is partially ordered by $\subseteq$ and $F \neq \varnothing$. Every chain Z in F has an upper bound $\bigcup Z$ in F. So by Zorn's lemma there exists a maximal element B in F. This B is an artificial attribute containing $\{a\}$.

COROLLARY Let $A = (D, \Delta, \mathrm{Con}, \vdash)$. If D is a denumerable infinite set and all artificial attributes in A are finite sets, then the number of artificial attributes in A is infinite.

Proof Assume now that the number of artificial attributes is finite. Since for any $a \in D$ there is an artificial attribute containing $\{a\}$, D is finite.

Let B_1, B_2 be artificial attributes in $A = (D, \Delta, \mathrm{Con}, \vdash)$. We say that

$$B_1 \leqslant B_2 \text{ iff } (\forall b \in B_2)(\forall c \in B_1)(c \subseteq b)$$

Let $\approx_B$, where B is an artificial attribute, be the relation on $\mathrm{Tot}_A \times \mathrm{Tot}_A$ defined as follows:

$$x \approx_B y \text{ iff } (\exists b \in B)(x \cap b = y \cap b = b)$$

THEOREM 3 If D is denumerable set and B is an artificial attribute that is finite, then $\approx_B$ is an equivalence relation on $\mathrm{Tot}_A \times \mathrm{Tot}_A$.

Proof The proof immediately follows from Theorem 1.

Assume now that $D_0 = \bigcup B$, where B is an artificial attribute, and let $\approx_{D_0}$ be the relation on Tot_A defined as follows:

$$x \approx_{D_0} y \text{ iff } x \cap D_0 = y \cap D_0$$

The question arises as to whether $\approx_B$ is equal to $\approx_{D_0}$.

FACT 5 If D is a denumerable set and $B = \{b_1, b_2, \ldots, b_n\}$ is an artificial attribute then $\approx_{D_0} \subseteq \approx_B$.

Proof Assume that $x \approx_{D_0} y$, which means that $x \cap D_0 = y \cap D_0$. From Theorem 1 we have that $\{\delta(b_i)\}_{i=1}^{n}$ is a partition of Tot_A. Hence $x \in \delta(b_i)$ for some $i \leqslant n$. Therefore $b_i \subseteq x$. Now, since $b_i = b_i \cap D_0 = b_i \cap D_0 \cap x = b_i \cap x$ and $b_i \cap D_0 \cap x = b_i \cap D_0 \cap y = b_i \cap y$, then $b_i \cap x = b_i \cap y = b_i$.

We will show now that the assumptions stated in Fact 5 are not strong enough to prove the inclusion $\approx_B \subseteq \approx_{D_0}$.

EXAMPLE Let us assume that $A = (D, \Delta, \mathrm{Con}, \vdash)$ where $D = \{r, a, m, u, g, y\}$ and

$$\mathrm{Con} = P(\{r, a, m\}) \cup P(\{r, a, u\}) \cup P(\{g, a, u\}) \cup P(\{r, y, m\})$$

By $P(C)$ we mean the power set of C. Let $B = \{\{r, a\}, \{g, a, u\}, \{r, y, m\}\}$. Clearly B is a maximal antichain,

$$\mathrm{Tot}_A = \{\{r, a, m\}, \{r, a, u\}, \{g, a, u\}, \{r, y, m\}\}$$

and

$$\{r, a, m\} \approx_B \{r, a, u\}$$

Taking now $D_0 = \{r, a, m, g, y\}$ we get $\{r, a, m\} \cap D_0 = \{r, a, m\} \neq \{r, a\} = \{r, a, u\} \cap D_0$, which means that $\sim(\{r, a, m\} \approx_{D_0} \{r, a, u\})$.

Let us assume from now that our information systems are with D denumerable and that all artificial attributes are finite sets.

FACT 6 Let B_1, B_2 be artificial attributes and $B_1 \leqslant B_2$. Then

- **(i)** If $b_1 \subseteq b_2$ and $b_1' \subseteq b_2$ then $b_1 = b_1'$ for any $b_1, b_1' \in B_1$, $b_2 \in B_2$
- **(ii)** $(\forall b_1 \in B)(\exists b_2 \in B_2)(b_1 \subseteq b_2)$
- **(iii)** $\approx_{B_2} \subseteq \approx_{B_1}$

Proof

- **(i)** Since $b_1 \subseteq b_2$ and $b_1' \subseteq b_2$, we have $b_1 \cup b_1' \subseteq b_2$. Now $b_2 \in \mathrm{Con}$, so $b_1 \cup b_1' \in \mathrm{Con}$. Hence $b_1 = b_1'$.
- **(ii)** Let $b_1 \in B_1$. Consider a total element $x \in \delta(b_1)$. Then $x \in \delta(b_2)$ for some b_2. From the definition of $B_1 \leqslant B_2$, we have $b_1' \subseteq b_2$ for some b_1'. So $b_1 \cup b_1' \subseteq x$, which implies $b_1 = b_1'$. Thus $b_1 \subseteq b_2$.
- **(iii)** Let $x \approx_{B_2} y$. By definition $x \supseteq b_2$ and $y \supseteq b_2$ for some b_2. From

$B_1 \leqslant B_2$, we have $b_1 \subseteq b_2$ for some $b_1 \in B_1$. So $x \supseteq b_1$ and $y \supseteq b_1$. Thus $x \approx_{B_1} y$.

Assume now that B_1 and B_2 are artificial attributes. Let

$$B_1 + B_2 \stackrel{\text{df}}{=} \{b_1 \cup b_2 : b_1 \in B_1,\ b_2 \in B_2,\ b_1 \cup b_2 \in \text{Con}\}$$

THEOREM 4 If B_1, B_2 are artificial attributes, then $B_1 + B_2$ is an artificial attribute and $\approx_{B_1+B_2} = \approx_{B_1} \cap \approx_{B_2}$.

Proof Let $b, b' \in B_1 + B_2$ and $b \neq b'$. Hence $b = b_1 \cup b_2$, and $b' = b_1' \cup b_2'$ where $b_1, b_1' \in B_1$, $b_2, b_2' \in B_2$. Since $b \neq b'$ then $b_1 \neq b_1'$ or $b_2 \neq b_2'$. Therefore

$$b \cup b' = (b_1' \cup b_1) \cup (b_2 \cup b_2') \notin \text{Con}$$

Now we have to prove that

$$\bigcup \{\delta(b) : b \in B_1 + B_2\} = \text{Tot}_A$$

Assume that $B_1 = \{b_1', b_2', \ldots, b_k'\}$, $B_2 = \{b_1^2, b_2^2, \ldots, b_n^2\}$. Since B_1, B_2 are artificial attributes, then

$$\bigcup_{i=1}^{k} \delta(b_i') = \text{Tot}_A \qquad \text{and} \qquad \bigcup_{j=1}^{n} \delta(b_j^2) = \text{Tot}_A$$

Hence,

$$\bigcup_{i=1}^{k} \delta(b_i') \cap \bigcup_{j=1}^{n} \delta(b_j^2) = \text{Tot}_A$$

From Fact 1 we have that

$$\bigcup_{i=1}^{k} \bigcup_{j=1}^{n} [\delta(b_i') \cap \delta(b_j^2)] = \bigcup_{i=1}^{k} \bigcup_{j=1}^{n} \delta(b_i' \cup b_j^2)$$

$$= \bigcup \{\delta(b_1 \cup b_2) : b_1 \in B_1,\ b_2 \in B_2\}$$

From Fact 2 we have that

$$\bigcup \{\delta(b_1 \cup b_2) : b_1 \in B_1,\ b_2 \in B_2\}$$
$$= \bigcup \{\delta(b_1 \cup b_2) : b_1 \in B_1,\ b_2 \in B_2,\ b_1 \cup b_2 \in \text{Con}\}$$
$$= \bigcup \{\delta(b) : b \in B_1 + B_2\}$$

FACT 7 If B_1, B_2, B_3 are artificial attributes, then $(B_1 + B_2) + B_3 = B_1 + (B_2 + B_3)$.

NOTATION If $B_1, B_2, \ldots, B_n$ are artificial attributes, then

$$\mathop{+}_{i=1}^{n} B_i \stackrel{\text{df}}{=} B_1 + \cdots + B_n$$

Rough Properties in Information Systems

In this section, we introduce the notion of a "testable attribute" and use it to build an approximation space [6]; then we introduce a first-order language to describe syntactically the semantic notion of rough properties.

Assume again that $A = (D, \Delta, \text{Con}, \vdash)$ is an information system with denumerable D, all artificial attributes being finite sets. We consider a situation when only a subset $D_1 \subseteq D$ of propositions (data objects) is "testable" ("visible"), which implies that we only work with rough information [2, 6] in the system. Let D_1 be a given subset of D called the set of "testable" propositions. An artificial attribute B is said to be testable iff $\bigcup B \subseteq D_1$. Clearly if $B_1, \ldots, B_n$ are all testable, then

$$\mathop{+}_{i=1}^{n} B_i$$

is testable. Let A be an information system and B be a testable artificial attribute. By an approximation space for A with respect to B we mean an ordered pair $(\text{Tot}_A, \approx_B)$.

Let L be the first-order language without constants and without functional letters. We denote by $z_1, z_2, \ldots$ its individual variables and by $r_i (i \in I)$ its predicate letters.

Now let us take an interpretation J of L in Tot_A and assume that $J(r_i) = R_i$ and n_i is the arity of R_i for any $i \in I$. Let B be a testable artificial attribute in A. For simplicity, we will write $\approx$ instead of $\approx_B$. Now we are ready to define two new interpretations of L in Tot_A called a lower $(\underline{J})$ and an upper $(\bar{J})$ approximation of J:

(i) $\underline{J}(r_i) = \underline{R}_i$ for any $i \in I$ where: $(x_1, x_2, \ldots, x_{ni}) \in \underline{R}_i$ iff for any $y_1, y_2, \ldots, y_{ni}$, if $y_1 \approx x_1, y_2 \approx x_2, \ldots, y_{ni} \approx x_{ni}$ then $(y_1, y_2, \ldots, y_{ni}) \in R_i$.

(ii) $\bar{J}(r_i) = \bar{R}_i$ for any $i \in I$ where $(x_1, x_2, \ldots, x_{ni}) \in \bar{R}_i$ iff there exist $y_1, y_2, \ldots, y_{ni}$ such that $y_1 \approx x_1, y_2 \approx x_2, \ldots, y_{ni} \approx x_{ni}$ and $(y_1, y_2, \ldots, y_{ni}) \in R_i$.

Let $\zeta(z_1, z_2, \ldots, z_n)$ be a formula in L, where $z_1, z_2, \ldots, z_n$ are all its free variables. By $J(\zeta)$ we mean the set

$$\{(x_1, x_2, \ldots) \in \text{Tot}_A^{\omega} : J \models \zeta[x_1, x_2, \ldots, x_n]\}$$

In a similar way we define $\underline{J}(\zeta)$ and $\bar{J}(\zeta)$.

THEOREM 5 If ζ is a positive (ie, without negation) formula in L, then $\underline{J}(\zeta) \subseteq J(\zeta) \subseteq \bar{J}(\zeta)$.

Proof By induction on the number of logical connectives in ζ,

(i) ζ is an atomic formula of the form $r_i(z_1, z_2, \ldots, z_n)$. Let $(x_1, x_2, \ldots) \in \underline{J}(\zeta)$. From the definition of $\underline{J}(\zeta)$ we have $(x_1, x_2, \ldots) \in \underline{J}(\zeta)$ iff $(y_1, y_2, \ldots) \in J(\zeta)$ for any $y_1, y_2, \ldots$ such that $x_1 \approx y_1, x_2 \approx y_2, \ldots, x_n \approx y_n$. Since $\approx$ is reflexive, then $(x_1, x_2, \ldots) \in J(\zeta)$. The proof that $J(\zeta) \subseteq \bar{J}(\zeta)$ is obvious.

(ii) ζ is of the form $\zeta_1 \wedge \zeta_2$. Let $(x_1, x_2, \ldots) \in \underline{J}(\zeta_1 \wedge \zeta_2)$. Since $\underline{J}(\zeta_1 \wedge \zeta_2) = \underline{J}(\zeta_1) \cap \underline{J}(\zeta_2)$, then $(x_1, x_2, \ldots) \in \underline{J}(\zeta_1)$ and $(x_1, x_2, \ldots) \in \underline{J}(\zeta_2)$. By induction hypothesis, $(x_1, x_2, \ldots) \in J(\zeta_1) \cap J(\zeta_2)$, where $J(\zeta_1) \cap J(\zeta_2) = J(\zeta_1 \wedge \zeta_2)$.

(iii) ζ is of the form $(\forall z)\psi(z, z_1, z_2, \ldots, z_n)$. Let $(x_1, x_2, \ldots) \in \underline{J}(\zeta)$. Hence, for any $x \in \mathrm{Tot}_A$, $(x, x_1, x_2, \ldots) \in \underline{J}(\psi)$. By induction hypothesis we have $(x, x_1, x_2, \ldots) \in J(\psi)$ for any $x \in \mathrm{Tot}_A$. Hence $(x_1, x_2, \ldots) \in J(\zeta)$.

The proof of the second inclusion is similar.

Corollary If ζ is a positive formula in L and $\underline{J} \models \zeta$ (ie, ζ is true in $\underline{J}$), then $J \models \zeta$).

The above corollary says that in order to verify the truth in J of any positive formula from L it is enough to verify its truth in $\underline{J}$.

Fact 8 If ζ is a formula in L such that $\underline{J}(\zeta) \subseteq J(\zeta) \subseteq \bar{J}(\zeta)$, then $\bar{J}(\sim\zeta) \subseteq J(\sim\zeta) \subseteq \underline{J}(\sim\zeta)$.

Proof Assume that $\underline{J}(\zeta) \subseteq J(\zeta) \subseteq \bar{J}(\zeta) \subseteq \mathrm{Tot}_A^\omega$. Hence, $\mathrm{Tot}_A^\omega - \bar{J}(\zeta) \subseteq \mathrm{Tot}_A^\omega - J(\zeta) \subseteq \mathrm{Tot}_A^\omega - \underline{J}(\zeta)$, which implies $\bar{J}(\sim\zeta) \subseteq J(\sim\zeta) \subseteq \underline{J}(\sim\zeta)$.

Now let us assume that $J(\zeta) = \approx$ for some $\psi(z_1, z_2)$ in L. This implies that $\underline{J}(\zeta)$ and $\bar{J}(\zeta)$, where ζ is from L, are definable in J by formulas from L. To prove it, we introduce two functions $\bar{T}$, $\underline{T}$ transforming formulas in L into formulas in L. The definition follows:

(i) If ζ is an atomic formula of the form $r_i(z_1, z_2, \ldots, z_{ni})$, then

$$\begin{aligned}\underline{T}(\zeta) &= \underline{T}(r_i(z_1, z_2, \ldots, z_{ni}))\\ &= (\forall z_1')(\forall z_2') \cdots (\forall z_{ni}')[\psi(z_1, z_1')\\ &\quad \wedge \cdots \wedge \psi(z_{ni}, z_{ni}') \rightarrow r_i(z_1', z_2', \ldots, z_{ni}')]\\ \bar{T}(\zeta) &= \bar{T}(r_i(z_1, z_2, \ldots, z_{ni}))\\ &= (\exists z_1')(\exists z_2') \cdots (\exists z_{ni}')[\psi(z_1, z_1')\\ &\quad \wedge \cdots \wedge \psi(z_{ni}, z_{ni}') \wedge r_i(z_1', z_2', \ldots, z_{ni}')]\end{aligned}$$

(ii) If ζ is of the form $\zeta_1 \wedge \zeta_2$, then

$$\bar{T}(\zeta) = \bar{T}(\zeta_1 \wedge \zeta_2) = \bar{T}(\zeta_1) \wedge \bar{T}(\zeta_2)$$

$$\underline{T}(\zeta) = \underline{T}(\zeta_1 \wedge \zeta_2) = \underline{T}(\zeta_1) \wedge \underline{T}(\zeta_2)$$

(iii) If ζ is of the form $\sim \zeta_1$, then

$$\underline{T}(\zeta) = \underline{T}(\sim \zeta_1) = \sim \underline{T}(\zeta_1)$$
$$\bar{T}(\zeta) = \bar{T}(\sim \zeta_1) = \sim \bar{T}(\zeta_1)$$

(iv) If ζ is of the form $(\forall z)\zeta_1$, then

$$\underline{T}(\zeta) = \underline{T}((\forall z)\zeta_1) = (\forall z)\underline{T}(\zeta_1)$$
$$\bar{T}(\zeta) = \bar{T}((\forall z)\zeta_1) = (\forall z)\bar{T}(\zeta_1)$$

It is easy to prove the following fact.

FACT 9 If ζ is a formula in L and $J(\zeta) = \approx$, then $\underline{J}(\zeta) = J(\underline{T}(\zeta))$ and $\bar{J}(\zeta) = J(\bar{T}(\zeta))$.

Proof The proof is by induction. We will show only that $\underline{J}(\zeta) = J(\underline{T}(\zeta))$ holds for atomic formulas. Assume $\zeta = r_i(z_1, z_2, \ldots, z_{ni})$ and $(x_1, x_2, \ldots) \in \underline{J}(\zeta)$. Hence $\underline{J} \models \zeta[x_1, x_2, \ldots, x_{ni}]$, which means that for any $y_1, y_2, \ldots, y_{ni}$, if $y_1 \approx x_1$, $y_2 \approx x_2, \ldots, y_{ni} \approx x_{ni}$ then $(y_1, y_2, \ldots, y_{ni}) \in R_i$. Therefore

$$J \models (\forall z_1')(\forall z_2') \cdots (\forall z_{ni}')[\Psi(z_1, z_1') \wedge \Psi(z_2, z_2') \wedge \cdots \wedge \Psi(z_{ni}, z_{ni}') \rightarrow r_i(z_1', z_2', \cdots, z_{ni}')[x_1, x_2, \cdots]]$$

which implies that $J \models \underline{T}(\zeta)[x_1, x_2, \ldots]$.

Let $S = (\mathrm{Tot}_A, \approx)$. By $\underline{q}$, where $q \subset \mathrm{Tot}_A$ we mean the union of all equivalence classes of the relation $\approx$ included in q, and by $\bar{q}$ we mean the least set containing q that is a union of equivalence classes of $\approx$. Sets $\underline{q}$ and $\bar{q}$ are called the lower and the upper approximation of q in S, respectively (see Pawlak [6]).

Let us take the extension L' of L by adding two functors $\bullet$, $-$ to its alphabet. In the standard interpretation, $\bullet$ means intersection and $-$ means complement. By T we denote the set of all terms in L'. Valuations are functions from the set $\{z_1, z_2, \ldots\}$ into $P(\mathrm{Tot}_A)$.

Let $t \in T$ be an arbitrary term in L' and v be a valuation. We define the value of t in the standard interpretation and under the valuation v denoted for simplicity by $t(v)$. We define also its lower approximation $\lfloor t(v) \rfloor$ and its upper approximation $\lceil t(v) \rceil$:

(i) $\lceil z_i(v) \rceil = \overline{v(z_i)}$
$\lfloor z_i(v) \rfloor = \underline{v(z_i)}$
$z_i(v) = v(z_i)$ for any $i \geqslant 1$

(ii) If t and t' are terms, then
$\lceil (t \bullet t')(v) \rceil = \lceil t(v) \rceil \cap \lceil t'(v) \rceil \qquad \lfloor (t \bullet t')(v) \rfloor = \lfloor t(v) \rfloor \cap \lfloor t'(v) \rfloor$

$$(t \bullet t')(v) = t(v) \cap t'(v) \qquad \lceil(-t)(v)\rceil = \mathrm{Tot}_A - \lfloor t(v)\rfloor$$
$$\lfloor(-t)(v)\rfloor = \mathrm{Tot}_A - \lceil t(v)\rceil \qquad (-t)(v) = \mathrm{Tot}_A - t(v)$$

THEOREM 6 For any $t \in T$ and a valuation v in Tot_A:
(i) $\lfloor t(v)\rfloor \subset t(v) \subset \lceil t(v)\rceil$
(ii) $\lfloor t(v)\rfloor \subset \underline{t(v)}$; $\overline{t(v)} \subset \lceil t(v)\rceil$.

Proof It is enough to prove only the first part of the theorem. The proof is by induction:
(i) $\lfloor z_i(v)\rfloor = \underline{v(z_i)} \subset v(z_i) \subset \overline{v(z_i)} = \lceil z_i(v)\rceil$
(ii) If t and t' are terms and

$$\lfloor t(v)\rfloor \subset t(v) \subset \lceil t(v)\rceil \qquad \lfloor t'(v)\rfloor \subset t'(v) \subset \lceil t'(v)\rceil$$

then $\lfloor(t \bullet t')(v)\rfloor = \lfloor t(v)\rfloor \cap \lfloor t'(v)\rfloor \subset t(v) \cap t'(v) \subset \lceil t'(v)\rceil \cap \lceil t(v)\rceil = \lceil(t \bullet t')(v)\rceil$.
(iii) If t is a term and $\lfloor t(v)\rfloor \subset t(v) \subset \lceil t(v)\rceil$, then $\mathrm{Tot}_A - \lceil t(v)\rceil \subset \mathrm{Tot}_A - t(v) \subset \mathrm{Tot}_A - \lfloor t(v)\rfloor$. Hence from the definition of $\lfloor t(v)\rfloor$ and $\lceil t(v)\rceil$, $\lfloor -t(v)\rfloor \subset -t(v) \subset \lceil -t(v)\rceil$.

THEOREM 7 If ζ is an open formula in L built from predicate letters $r_1, \ldots, r_k$ and J is an interpretation of L, then

$$\lfloor J(\zeta)\rfloor \subset J(\zeta) \subset \lceil J(\zeta)\rceil$$

where $\lfloor J(\zeta)\rfloor$ and $\lceil J(\zeta)\rceil$ are defined as follows:
(i) $\lfloor J(r_i(z_1, \ldots, z_{ni}))\rfloor = \{(x_1, x_2, \ldots) \in \mathrm{Tot}_A^\omega : (x_1, \ldots, x_{ni}) \in \underline{R}_i\}$
$\lceil J(r_i(z_1, \ldots, z_{ni}))\rceil = \{(x_1, x_2, \ldots) \in \mathrm{Tot}_A^\omega : (x_1, \ldots, x_{ni}) \in \overline{R}_i\}$
(ii) $\lfloor J(\zeta_1 \wedge \zeta_2)\rfloor = \lfloor J(\zeta_1)\rfloor \cap \lfloor J(\zeta_2)\rfloor$; $\lceil J(\zeta_1 \wedge \zeta_2)\rceil = \lceil J(\zeta_1)\rceil \cap \lceil J(\zeta_2)\rceil$
(iii) $\lfloor J(\sim \zeta_1)\rfloor = \mathrm{Tot}_A - \lceil J(\zeta_1)\rceil$; $\lceil J(\sim \zeta_1)\rceil = \mathrm{Tot}_A - \lfloor J(\zeta_1)\rfloor$.

CONCLUSION

The notion of an artificial attribute in Scott's system enables us to define an approximation space that next becomes the domain for interpretations of the first-order language. Scott's system gives us the opportunity to talk about objects as infinite consistent sets of propositions. Knowing only a finite consistent set of propositions, we may talk in a very natural way about incomplete information. We plan to extend our first-order language L by adding to it propositions from Scott's system and in that way to have a better link between language L and Scott's systems. The main reason for introducing the first-order language here is to have an opportunity to talk in some formalized language about rough sets of elements in Scott's system, rough relations among them, and their rough properties.

References

1. Salton, G., *Automatic Information Organization and Retrieval,* McGraw-Hill, New York, 1968.
2. Pawlak, Z., Rough classification, *Int. J. Man-Mach. Stud.,* 20, 469–483, 1984.
3. Scott, D., Domains for denotational semantics, a corrected and expanded version of a paper prepared for ICALP'82, Aarhus, Denmark, 1982.
4. Ras, Z., An algebraic approach to information retrieval systems, *Int. J. Comp. Inf. Sci.,* 11(4), 275–293, 1982.
5. Pawlak, Z., Classification of objects by means of attributes, ICS PAS Reports, No. 429, IPI PAN, Warsaw, Poland, 1981.
6. Pawlak, Z., Rough sets, *Int. J. Inform. Comp. Sci.* 11, 341-356, 1982.